Exhibit C

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 277, No. 38, Issue of September 20, pp. 35574–35585, 2002
Printed in U.S.A.

Identification, Characterization, and Functional Study of the Two Novel Human Members of the Semaphorin Gene Family*Ⓢ

Received for publication, June 28, 2002
Published, JBC Papers in Press, July 10, 2002, DOI 10.1074/jbc.M206451200

Xianghu Qu‡, Handong Wei‡, Yun Zhai‡, Haiping Que§, Qian Chen§, Fei Tang‡, Yan Wu§, Guichun Xing‡, Yunping Zhu‡, Shaojun Liu§, Ming Fan§, and Fuchu He‡¶

From the ‡Department of Genomics and Proteomics, Beijing Institute of Radiation Medicine, Chinese National Human Genome Center at Beijing and the §Department of Neuroscience, Beijing Institute of Basic Medical Sciences, Beijing 100850, China

We cloned two novel human transmembrane semaphorins, (HSA)SEMA6C and (HSA)SEMA6D, that belong to the class VI subgroup of the semaphorin family. The genes for SEMA6C and SEMA6D are mapped on chromosome 1q12–21.1 and 15q21.1, respectively. Among the adult tissues, SEMA6C is expressed only in skeletal muscle, whereas SEMA6D is expressed abundantly in kidney, brain, and placenta and moderately in the heart and skeletal muscles. During murine development, neither SEMA6C nor SEMA6D was expressed in embryonic day 10.5 (E10.5) embryos, but both were highly expressed in the areas of the lateral ventricle, the striatum, the wall of the midbrain, the pons/midbrain junction, and the choroid plexus of E13 embryos. Were neurons, neither axons nor astrocytes, highly expressed both semaphorins. Three isoforms of SEMA6C and five isoforms of SEMA6D derived from alternative splicing were identified, and their expression was regulated in a tissue- and development-dependent manner. Deletion analysis indicated that a sema domain and a PSI domain are integrally necessary for correct post-translation modification and subcellular localization. The extracellular domain of SEMA6C inhibited axonal extension of nerve growth factor-differentiated PC12 cells and induced the growth cone collapse of chicken dorsal root ganglion, rat hippocampal neurons, and rat cortical neurons in a dose-responsive manner. SEMA6D acted like SEMA6C except it had no significant effect on the growth cones of rat cortical neurons.

Growing axons navigate through the developing embryo with remarkable accuracy. An axon's response to the guidance cues in its immediate environment determines the trajectory it will take. Within the past few years, many guidance cues belonging to the semaphorin (1–3), netrin (4), ephrin (5), and slit (6, 7) signaling molecules families, and to Nogo (8) of the Reticulon family, have all been shown to attract or repel specific axons in culture and/or to affect axon guidance *in vivo*.

All semaphorins contain a semaphorin (sema)[1] domain and a PSI domain (found in plexins, semaphorins, and integrins) (9) in the extracellular portion and a class-specific C terminus that may contain additional sequence motifs. At present, semaphorins have been categorized into eight subclasses. Class I and class II semaphorins are found in invertebrates. Classes III to VII are found in vertebrates, and a final class comprises proteins encoded by viruses. In vertebrates, class III semaphorins are secreted proteins, whereas classes IV–VI are transmembrane proteins. Class VI and class III semaphorins resemble class I and class II semaphorins, respectively, in their domain arrangements, but both class VI and class III semaphorins are phylogenetically distinct from class I and class II semaphorins (1–3).

Recently, many new genes with important biological functions on cell migration have been detected based on large scale sequencing of human fetal liver cDNA libraries in our laboratory (10, 11). To identify potential novel human semaphorin genes, more than 14,000 expressed sequence tags (ESTs) of human fetal liver cDNA libraries sequenced in our laboratory (12) have been analyzed to search the non-redundant GenBank™ data base with the Blast program. Among these ESTs, an insert clone, FLD6219, with high homology to rat *Sema6C* (R-Sema Y) cDNA (13) was selected for further study.

Here, we report on the molecular cloning, mapping, and functional analysis of two mammalian semaphorins, (HSA)SEMA6C and (HSA)SEMA6D. Sequencing of the two genes for SEMA6C and SEMA6D indicated that both are class VI transmembrane semaphorins. Three isoforms of SEMA6C and five isoforms of SEMA6D, probably generated by alternative splicing, were identified. Functional studies show that SEMA6C and SEMA6D not only induce growth cone collapse of dorsal root ganglion (DRG) and certain cultured rat neurons but also inhibit axonal extension of nerve growth factor (NGF)-differentiated PC12 cells in a dose-responsive manner. The expression profiles and genomic organization of the genes for SEMA6C and SEMA6D are also described.

* This work was supported in part by the Chinese National Key Project of Basic Research, the Chinese High-tech Program, and Grants 30070177 and 30070285 from the Chinese National Natural Sciences Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Ⓢ The on-line version of this article (available at *http://www.jbc.org*) contains supplemental data.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF339154 (human SEMA6C.1), AF339152 (human SEMA6C.2), AF339153 (human SEMA6C.3), AF363973 (mouse Sema6C.1), AF363972 (mouse Sema6C.2), AF363971 (rat Sema6C.2), AF389430 (SEMA6D.1), AF389427 (SEMA6D.2), AF389428 (SEMA6D.3), AF389429 (SEMA6D.4), and AF389426 (SEMA6Ds).

¶ To whom correspondence should be addressed: Dept. of Genomics and Proteomics, Beijing Institute of Radiation Medicine, 27 Taiping Rd., Beijing 100850, China. Tel.: 8610-68171208; Fax: 8610-68214653, E-mail: hefc@nic.bmi.ac.cn.

[1] The abbreviations used are: sema, semaphorpin; PSI, found in plexins, semaphorins, and integrins; EST, expressed sequence tag; RT-PCR, reverse transcription PCR; CDS, coding sequence; RACE, rapid amplification of cDNA ends; DRG, dorsal root ganglion; NGF, nerve growth factor; aa, amino acid(s); contig, contiguous group of overlapping clones; DEPC, diethyl pyrocarbonate; PBS, phosphate-buffered saline; MOPS, 4-morpholinepropanesulfonic acid; E, embryonic day (*e.g.* E10.5); NBT, nitro blue tetrazolium; BCIP, 5-bromo-4-chloro-3-indolyl phosphate.

This paper is available on line at http://www.jbc.org

EXPERIMENTAL PROCEDURES

Identification and cDNA Cloning of Human SEMA6C and SEMA6D—Based on a cDNA clone (No. FLD6219) homologous to rat *Sema6C* cDNA, we obtained a putative full-length cDNA of SEMA6C using contig from public data bases, PCR with primers directly based on the rat *Sema6C* sequence, and the 5′-RACE (rapid amplification of cDNA ends) technique (SMART™ RACE cDNA Amplification Kit, CLONTECH). Then, two sets of primers were used to confirm existence of the predicted cDNA sequence of SEMA6C and to amplify its coding sequence (CDS) from human brain cDNA. These resulted in pGEM-T vectors with the entire CDSs of the three alternative splicing variants (SEMA6C.1, SEMA6C.2, and SEMA6C.3), respectively. To identify any SEMA6C-related semaphorins in human, we performed a contig search through sequence databases with SEMA6C.1 cDNA and retrieved two partial cDNAs (5223 and 1588 bp) susceptible to encoding a SEMA6C-related semaphorin. Then reverse transcription (RT)-PCRs and 5′-RACE were used to isolate the full-length cDNA. These resulted in pGEM-T vectors with the entire CDSs of the five alternative splicing variants (SEMA6D.1, SEMA6D.2, SEMA6D.3, SEMA6D.4, and SEMA6Ds), respectively. The primers used above are described in the Supplemental Material.

To distinguish three isoforms of SEMA6C or four long isoforms of SEMA6D, RT-PCRs were performed using poly(A)$^+$ RNAs of human tissues (CLONTECH) and total RNAs of rat and mouse tissues. A pair of primers for detecting the 96-bp insertion in *SEMA6C* were as follows: P7, 5′-AACAGCACGACGGATCATAG-3′; and P8, 5′-GAGGAGTGG-GATGGGGAC-3′. A pair of primers for detecting the 120-bp insertion in *SEMA6C* were as follows: P9, 5′-GTCTGCGCCTTCTACCTGGA-3′; and P10, 5′-GGCCTGTAAAACATCAAAT-3′. A pair of primers for detecting the long isoforms of SEMA6D were as follows: P11, 5′-TAGAGTGAC-CCCAGGGATGC-3′; and P12, 5′-GACTGGACTTCCCATCGTAC-3′. The PCR amplification for *SEMA6C* was performed under the following conditions: 4 min of initial denaturation at 94 °C followed by 30 cycles of 94 °C for 45 s, 58 °C for 45 s, and 72 °C for 1 min and ending with a final extension at 72 °C for 7 min. Amplification used TaKaRa La *Taq* with GC buffer (TaKaRa Biotechnology Inc., Dalian, China). The PCR program for *SEMA6D* comprised 94 °C for 4 min followed by 30 cycles of 94 °C for 45 s, 55 °C for 45 s, and 72 °C for 45 s. PCR products were separated on agarose gels (1% for *SEMA6C* and 2% for *SEMA6D*), purified from the gel, and sequenced directly. The principle of computer-based chromosomal assignment of a new gene was described previously (11).

Sequence Analysis of Human SEMA6C and SEMA6D—The sequence alignment and phylogenetic tree among mammalian class semaphorin proteins were performed using the program CLUSTAL W. Protein subsequence motifs were identified using the network service SMART (smart.embl-heidelberg.de). Prosite information of the proteins was analyzed with the TMpred program, which was used to make a prediction of membrane-spanning regions and their orientation.

Northern Blot Analysis—A human multiple tissue Northern blot (CLONTECH) was hybridized to the specific probes of SEMA6C, SEMA6D, and actin, respectively. The probes were generated with the Klenow fragment of DNA polymerase I and [α-^{32}P]dCTP using the Prime-a-Gene® labeling system (Promega). The Northern blot was premade with Poly(A)$^+$ RNA from 12 human tissues. The DNA templates for the probes of SEMA6C (an *Xho*I-*Sac*I fragment of 871 bp, nucleotides 2820–3691 of SEMA6C.1), the SEMA6D long isoform-specific probe (a *Bam*HI fragment of 926 bp, nucleotides 2055–2981 of SEMA6D.1), the SEMA6D-specific probe (nucleotides 1875–2166 of SEMA6Ds), and the SEMA6D common probe (a *Sac*I-*Pst*I fragment of 538 bp, nucleotides 216–853 of SEMA6D.1) were the fragments that were not homologous to other members of this family. The hybridization procedure followed the manufacturer's protocol. After being washed, the blots were exposed to x-ray film at −70 °C with an intensifying screen.

Expression Vectors and Immunofluorescence Assay—To express semaphorin extracellular domains in soluble forms, the extracellular domain (including sema domain and PSI domain) of SEMA6C.1 (aa 1–573) or SEMA6C.3 (aa 1–533) cDNA was subcloned into Myc/His-tagged expression vector pcDNA3.1-MycHis (Invitrogen) at the *Kpn*I/*Eco*RI sites to create the constructs allowing the expression of C terminally tagged Myc-His fusion proteins, *i.e.* SEMA6C.1-mh and SEMA6C.3-mh, respectively. The constructed plasmids are named pcDNA3.1-SEMA6C.1 and pcDNA3.1-SEMA6C.3, respectively. Similarly, the extracellular domain of SEMA6D.1 (aa 1–592) cDNA and the entire CDS of SEMA6Ds were subcloned in-frame into the pcDNA3.1 vector at the *Not*I/*Xba*I sites and *Bam*HI/*Eco*RI sites, leading to pcDNA3.1-SEMA6D.1 and pcDNA3.1-SEMA6Ds, respectively. pcDNA3.1-SEMA6C.1ΔPSI was constructed by deleting the PSI domain from the pcDNA3.1-SEMA6C.1 construct. In brief, two fragments were amplified from the plasmid pcDNA3.1-SEMA6C.1 using two sets of primers and were then cut and religated to pcDNA3.1-MycHis. The following primers were used: P13 (sense), CGGGGTACCATGCCCCGTGC-CCCCCACTTCATGC, and P14 (antisense), 5′-CGGGTCTAGAAAAAGC-CACAAAAAGCCTGTG-3′; P15 (sense), 5′-CGGGTCTAGAGCTACTGG-GAGTCAGTCTGGC-3′, and P16 (antisense), 5′-CGGAATTCAAAGTTG-AAACGGCCGCCGTTCGGG-3′. The subcloned inserts were confirmed by sequencing.

For intracellular localization study of the five fusion proteins mentioned above, COS7 cells transfected with pcDNA3.1-SEMA6C.1, pcDNA3.1-SEMA6C.3, pcDNA3.1-SEMA6C.1-ΔPSI, pcDNA3.1-SEMA6D.1, or pcDNA3.1-SEMA6Ds were fixed in 30% paraformaldehyde, permeabilized in 0.5% Triton X-100, and stained with anti-Myc antibody according to standard procedure (14). The nuclei were stained by Hoechst 33342. All cell samples were viewed under an epifluorescence microscope.

Isolation and Culture of Primary Neurons and in Situ Hybridization—The primary neurons were isolated from the spinal cords of newborn rats (12 h). The spinal cords were dissected out and cut with scissors into about 1-mm^3 blocks. Dulbecco's modified Eagle's medium with 0.25% trypsin was added and digested for 30 min at 37 °C, which was terminated by adding the basic culture medium (Dulbecco's modified Eagle's medium with 10% fetal bovine serum and 1% N3). After 1 min, the tissue was transferred into basic medium and rinsed twice. The tissue blocks were then blown gently with a Pasteur pipette, and then the isolated neurons were cultured in a basic medium at a density of 0.5×10^5/cm^2. After 1 day, arabinosylcytosine (10^{-5} M) was added into the medium, which was then cultured continuously for another day to kill dividing cells. Every day thereafter, half of the medium was replaced with appropriate fresh medium. After removing arabinosylcytosine throughout by two changes of the N3 culture medium, the neurons were cultured continuously for another day, and then the neurons were fixed with in 4% paraformaldehyde and 1/1000 DEPC solution (in PBS, pH 7.4) at room temperature for 30 min. Then, the nonradioactive *in situ* hybridization protocols were employed as follows. The neurons were dipped in 0.5% methanol to extinguish endogenesis peroxidase for 30 min at room temperature. After being washed, the culture dishes were incubated in a pepsin solution for 1 min at room temperature to expose the mRNA section, and then 20 μl of pre-hybridization solution was added to the dishes for 4 h at 37 °C. Afterward, 20 μl of hybridization solution including differently labeled probes were added separately into the dishes overnight at 40 °C. After being washed thoroughly with SSC, the dishes were incubated in antibody against digoxigenin, conjuncted by biotin for 60 min at 37 °C, and visualized with the BCIP/NBT system for 20–30 min at room temperature.

Whole Mount in Situ Hybridization—Murine embryos (10.5 and 13 days) were dissected free of any extra-embryonic membranes, fixed in 4% paraformaldehyde in DEPC-PBS overnight at 4 °C, washed in DEPC-PBS, and then stored in 100% methanol at −20 °C until required. The embryos were rehydrated in gradient methanol, washed twice in PTw (1× PBS, 0.1% Tween 20) for 5 min each, treated with 10 μg/ml proteinase K in PTw for 30 min, rinsed once gently in PTw, re-fixed for 20 min at room temperature, washed twice in PTw for 5 min at room temperature, and transferred to an 0.5-ml Eppendorf tube. Then as much liquid as possible was removed while taking care to avoid damaging the embryos; 0.5 ml hybridization mix (50% formamide, 5× SSC, 0.5 mg/ml yeast tRNA, 0.5 mg/ml salmon sperm DNA, 5× Denhardt's solution, 0.5% SDS) was added. The mixture was removed and replaced with fresh hybridization mix, the embryos were pre-hybridized at 63 °C for 2 h, probe was added to a final concentration of about 0.5 μg/ml; the mixture was hybridized overnight at 63 °C and was then rinsed twice with pre-warmed 63 °C hybridization mix, digested with 20 μg/ml RNase A (10 mmol/liter Tris-HCl, 0.5 mol/liter NaCl, 5 mmol/liter EDTA, pH 8.0) for 30 min, twice for 20 min in 2× SSC at 63 °C, twice for 20 min in 0.1× SSC at 63 °C, and once in buffer 1 (0.1 mol/liter maleic acid, 0.15 mol/liter NaCl, pH 7.5). After a 60-min blocking at room temperature in 5% normal goat serum in buffer 2 (1% blocking reagent in buffer 1; Roche Molecular Biochemicals), embryos were incubated overnight at 4 °C in a 1:2000 dilution of anti-digoxigenin Fab fragment in 5% normal goat serum in buffer 2. They were washed four times with buffer 1 for 1 h each, washed twice with alkaline phosphatase buffer (0.1 mmol/liter Tris-HCl, pH 9.5, 0.1 mmol/liter NaCl, and 50 mmol/liter $MgCl_2$) for 1 h each. For every ml of alkaline phosphatase buffer, 4.5 μl of NBT and 3.5 μl of BCIP were added, and the mixture was developed in the dark for 2 to 20 h. When the reaction had pro-

```
Sema6C_MMU   MPRAPHSMPLLLLLLLLSSLPQAQAAFPQDPTPLLTSDLQGASPSSWFRGLEDDAVAAEL 60
Sema6C_RNO   MPRAPHSMPLLLLLLLLS--LPQAQTAFPQDPIPLLTSDLQGTSPSSWFRGLEDDAVAAEL 58
SEMA6C_HSA   MPRAPHFMPLLLLLLLLS-LPHTQAAFPQDPLPLLISDLQGTSPLSWFRGLEDDAVAAEL 59
SEMA6D_HSA   -MRVFLLCAYILLLMVS---QLRAVSFPEDDEPLNTVDYHYSRQYPVFRGRPSGNESQHR 56
              *.    . :***::          .:**:*  **   * : :    . ***  ..  : .

Sema6C_MMU   GLDFQRFLTLNRTLLVAARDHVFSFDLQAQEEGEGLVPNKFLTWRS--QDMENCAVRGKL 118
Sema6C_RNO   GLDFQRFLTLNRTLLVAARDHVFSFDLQAQEEGEGLVPNKFLTWRS--QDMENCAVRGKL 116
SEMA6C_HSA   GLDFQRFLTLNRTLLVAARDHVFSFDLQAEEEGEGLVPNKYLTWRS--QDVENCAVRGKL 117
SEMA6D_HSA   -LDFQLMLKIRDTLYIAGRDQVYTVNLNEMPKTE-VIPNKKLTWRSRQQDRENCAMKGKH 114
              **** :*.:. ** :*.**:*::.:*:    : * ::*** *****  ** ****::**

Sema6C_MMU   TDECYNYIRVLVPWNSQTLLACGTNSFSPMCRSYGITSLQQEGEELSGQARCPFDATQST 178
Sema6C_RNO   TDECYNYIRVLVPWDSQTLLACGTNSFSPVCRSYGITSLQQEGEELSGQARCPFDATQST 176
SEMA6C_HSA   TDECYNYIRVLVPWDSQTLLACGTNSFSPVCRSYGITSLQQEGEELSGQARCPFDATQSN 177
SEMA6D_HSA   KDECHNFIKVFVPRNDEMVFVCGTNAFNPMCRYYRLSTLEYDGEEISGLARCPFDARQTN 174
             .***:*:*:*:** :.: ::.****:*.*:** * :::*: :***:** ******* *:.

Sema6C_MMU   VAIFAEGSLYSATAADFQASDAVVYRSLGPQPPLRSAKYDSKWLREPHFVYALEHGEHVY 238
Sema6C_RNO   VAISAEGSLYSATAADFQASDAVVYRSLGPQPPLRSAKYDSKWLREPHFVYALEHGDHVY 236
SEMA6C_HSA   VAIFAEGSLYSATAADFQASDAVVYRSLGPQPPLRSAKYDSKWLREPHFVQALEHGDHVY 237
SEMA6D_HSA   VALFADGKLYSATVADFLASDAVIYRSMGDGSALRTIKYDSKWIKEPHFLHAIEYGNYVY 234
             **: *:*.*****.*** *****:***:*  ..**: ******::****: *:*:*::**

Sema6C_MMU   FFFREVSVEDARLGRVQFSRVARVCKRDMGGSPRALDRHWTSFLKLRLNCSVPGDSTFYF 298
Sema6C_RNO   FFFREVSVEDARLGRVQFSRVARVCKRDMGGSPRALDRHWTSFLKLRLNCSVPGDSTFYF 296
SEMA6C_HSA   FFFREVSVEDARLGRVQFSRVARVCKRDMGGSPRALDRHWTSFLKLRLNCSVPGDSTFYF 297
SEMA6D_HSA   FFFREIAVEHNNLGKAVYSRVARICKNDMGGSQRVLEKHWTSFLKARLNCSVPGDSFFYF 294
             *****::**.  .**:. :*****:**.***** *.*::******* ********** ***

Sema6C_MMU   DVLQSLTGPVNLHGRSALFGVFTTQTNSIPGSAVCAFYLDDIERGFEGKFKEQRSLDGAW 358
Sema6C_RNO   DVLQSLTGPVNLHGRSALFGVFTTQTNSIPGSAVCAFYLDDIERGFEGKFKEQRSLDGAW 356
SEMA6C_HSA   DVLQALTGPVNLHGRSALFGVFTTQTNSIPGSAVCAFYLDEIERGFEGKFKEQRSLDGAW 357
SEMA6D_HSA   DVLQSITDIIQINGIPTVVGVFTTQLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTPDSVW 354
             ****::*. ::::* .::.****** ***********  :*:**: *:*:****:: *..*

Sema6C_MMU   TPVSEDKVPSPRPGSCAGVGAAASFSSSQDLPDDVLLFIKAHPLLDPAVPPATHQPLLTL 418
Sema6C_RNO   TPVSEDKVPSPRPGSCAGVGAAALFSSSQDLPDDVLLFIKAHPLLDPAVPPATHQPLLTL 416
SEMA6C_HSA   TPVSEDRVPSPRPGSCAGVGGAALFSSSRDLPDDVLTFIKAHPLLDPAVPPVTHQPLLTL 417
SEMA6D_HSA   TAVPEDKVPKPRPGCCAKHGLAEAYKTSIDFPDETLSFIKSHPLMDSAVPPIADEPWFTK 414
             *.*.**:**.****.**  *  *  :.:* *:**:.* ***:***:*.****  :.* :*
```

FIG. 1. **Amino acid sequence alignment of (HSA)SEMA6D.1, (HSA)SEMA6C.1, rat Sema6C, and mouse Sema6C.** The most conserved region is located in their extracellular domains, *i.e.* the Sema domain and the PSI domain. The six conserved cysteines in the cysteine-rich motif, $CX_{(8)}CX_{(5)}CX_{(3)}CX_{(7)}CX_{(7/8)}C$, are shown.

```
Sema6C_MMU    T-SRALLTQVAVDGMAGPHRNTTVLFLGSNDGTVLKVLPPGGQSLGSEPIVLEEIDAYSH 477
Sema6C_RNO    T-SRALLTQVAVDGMAGPHRNTTVLFLGSNDGTVLKVLPPGGQSLGPEPIILEEIDAYSH 475
SEMA6C_HSA    T-SRALLTQVAVDGMAGPHSNITVMFLGSNDGTVLKVLPPGGRSGGPEPILLEEIDAYSP 476
SEMA6D_HSA    TRVRYRLTAISVDHSAGPYQNYTVIFVGSEAGMVLKVLAKTSPFSLNDSVLLEEIEAYNH 474
              *  *  ** ::**  ***: * **:*:**: * *****.  .     :.::****:**.

Sema6C_MMU    ARCSGKRSPRAARRIIGLELDTEGHRLFVAFPGCIVYLSLSRCARHGACQRSCLASLDPY 537
Sema6C_RNO    ARCSGKRSPRAARRIIGLELDTEGHRLFVAFPGCIVYLSLSRCARHGACQRSCLASLDPY 535
SEMA6C_HSA    ARCSGKRTAQTARRIIGLELDTEGHRLFVAFSGCIVYLPLSRCARHGACQRSCLASQDPY 536
SEMA6D_HSA    AKCSAEN--EEDKKVISLQLDKDHHALYVAFSSCIIRIPLSRCERYGSCKKSCIASRDPY 532
              *:**.:.    . :::*.*:**.: * *:***..**: :.**** *:*:*::**:** ***

Sema6C_MMU    CGWHRSRGCMSIR-GPGGTDVDLTGNQESTEHGDCQDGATGSQSGPGDSAYVLLGPGRSP 596
Sema6C_RNO    CGWHRFRGCVNIR-GPGGTDVDLTGNQESMEHGDCQDGATGSQSGPGDSAYVLLGPGPSP 594
SEMA6C_HSA    CGWHSSRGCVDIR-GSGGTDVDQAGNQESMEHGDCQDGATGSQSGPGDSAYVLPGPGPSP 595
SEMA6D_HSA    CGWLSQGSCGRVTPGMLLLTEDFFAFHNHSAEGYEQDTEFGNTAHLGD------------ 580
              ***    .*  :  *       *  . ::   .*   **   *. :  **

Sema6C_MMU    ETPSSPSDAQQGPQSSTLGAHTQGVRRDLSPASASRSIPIPLLLACVAAAFALGASVSGL 656
Sema6C_RNO    ETPSSPSDAHPGPQSSTLGAHTQGVRRDLSPASASRSIPIPLLLACVAAAFALGASVSGL 654
SEMA6C_HSA    GTPSPPSDAHPRPQSSTLGVHTRGVRRDLPPASASRSVPIPLLLASVAAAFALGASVSGL 655
SEMA6D_HSA    ---------------------CHGVRWEVQSGESNQMVHMNVLITCVFAAFVLGAFIAGV 619
                                   :*** :: ...:.: : : :*::.* ***.*** ::*:

Sema6C_MMU    LVSCA---CRRANRRRSKDIETPGLPRPLSLRSLARLHGGGPEPPPPPKDGDAAQTPQLY 713
Sema6C_RNO    LVSCA---CRRANRRRSKDIETPGLPRPLSLRSLARLHGGGPEPPPPPKDGDAAQTPQLY 711
SEMA6C_HSA    LVSCA---CRRAHRRRGKDIETPGLPRPLSLRSLARLHGGGPEPPPPSKDGDAVQTPQLY 712
SEMA6D_HSA    AVYCYRDMFVRKNRKKIHKDAES-AQSCTDSSGSFAKLNG--LFDSPVKEYQQNIDSPKLY 676
               * *      * :*: ** *: . . . * *:*:*:*      .*   :  :  ::*:**

Sema6C_MMU    TTFLP-----PPDG----------GSPPELACLPTPETTPELPVKHLRASGGPWEWNQ-N 757
Sema6C_RNO    TTFLP-----PPEG----------GSPPELACLPTPETTPELPVKHLRASGGPWEWNQ-N 755
SEMA6C_HSA    TTFLP-----PPEG----------VPPPELACLPTPESTPELPVKHLRAAGDPWEWNQ-N 756
SEMA6D_HSA    SNLLTSRKELPPNGDTKSMVMDHRGQPPELAALPTPESTPVLHQKTLQAMKSHSEKAHGH 736
              :.:*.     **:*              *****.*****:** *   * *:*  .   *  : :

Sema6C_MMU    GNNASEG----PGRPPRGCSGAGGPAPRVLVRP--------------------------- 786
Sema6C_RNO    GNNASEG----PGRP-RGCSAAGGPAPRVLVRP--------------------------- 783
SEMA6C_HSA    RNNAKEG----PGRS-RGGHAAGGPAPRVLVRP--------------------------- 784
SEMA6D_HSA    GASRKETPQFFPSSPPPHSPLSHGHIPSAIVLPNATHDYNTSFSNSNAHKAEKKLQNIDH 796
                . .*     *. .        : *  * .:* *
```

FIG. 1—*continued*

```
Sema6C_MMU     PPPGCPGQ-----AVEVT-TLEELLRYLHGPQPP--------RKGSEPLASAPFTSR---  829
Sema6C_RNO     PPPGCPGQ-----EVEVT-TLEELLRYLHGPQPP--------RKGSEPLASAPFTSR---  826
SEMA6C_HSA     PPPGCPGQ-----AVEVT-TLEELLRYLHGPQPP--------RKGAEP--PAPLTSR---  825
SEMA6D_HSA     PLTKSSSKRDHRRSVDSRNTLNDLLKHLNDPNSNPKAIMGDIQMAHQNLMLDPMGSMSEV  856
               *  , ...:     *:   **::**::*:.*:.          : . :    *: *

Sema6C_MMU     -PPASEPGASLFVDSSPMPRDGVPPLRLDVPPE------------GKRAAPSGRPALSAP  876
Sema6C_RNO     -PPASEPGAALFVDSSPMPRDCVPPLRLDVPPD------------GKRAAPSGRPALSAP  873
SEMA6C_HSA     -ALPPEPAPALLGGPSPRPHECASPLRLDVPPE------------GRCASAPARPALSAP  872
SEMA6D_HSA     PPKVPNREASLYSPPSTLPRN-SPTKRVDVPTTPGVPMTSLERQRGYHKNSSQRHSISAM  915
                .   .:  .:*   .*.  *::  ..  *:***.             *   ..  * ::**

Sema6C_MMU     APRLGVGGSRRLPFPTHRAPPGLLTRVPSGGPARYSGGP--GRHLLY-LGRPEGHRGR--  931
Sema6C_RNO     APRLGVSGSRRLPFPTHRAPPGLLTRVPSGGPSRYSGGP--GRHLLY-LGRPDGHRGR--  928
SEMA6C_HSA     APRLGVGGGRRLPFSGHRAPPALLTRVPSGGPSRYSGGP--GKHLLY-LGRPEGYRGR--  927
SEMA6D_HSA     PKNLNSPNGVLLSRQPSMNRGGYMP-TPTGAKVDYIQGTPVSVHLQPSLSRQSSYTSNGT  974
               . .*.  ..   *.      . *:.  .*:*.    *  *.   . **   *.* ..: ..

Sema6C_MMU     ----SLKRVDVKSP-LSPKPPLASPP--QPAPHGGHFNF  963
Sema6C_RNO     ----SLKRVDVKSP-LSPKPPLATPP--QPAPHGSHFNF  960
SEMA6C_HSA     ----ALKRVDVEKPQLSLKPPLVGPSSRQAVPNGGRFNF  962
SEMA6D_HSA     LPRTGLKRTPSLKPDVPPKPSFVPQT--PSVRPLNKYTY  1011
                   .***.   .* :.  **.:.  .   ..   .::.:
```

FIG. 1—*continued*

ceeded to our satisfaction, it was imperative to quickly stop the reaction to prevent excess background. Embryos were washed twice in alkaline phosphatase buffer for 5 min each and then washed at least three times in PTw buffered to pH 5.5 for 1 h each in the dark and fixed in MEMFA (0.1 M MOPS, pH 7.5, 2 mM EGTA, 1 mM $MgSO_4$, 3.7% formaldehyde) for 1 h. After this time, the embryos could be cleared in glycerol for visualization.

Growth Cone Collapse Assay and Neurite Outgrowth Assay—COS7 cells (5×10^5) were transfected with 4 μg of pcDNA3.1-MycHis vector, pcDNA3.1-SEMA6C.1, or pcDNA3.1-SEMA6D.1 using LipofectAMINE (Invitrogen) as recommended by the manufacturer's protocol. 36–48 h after the transfection, the cells and the conditioned media were collected as described by Luo *et al.* (15). Serum-free media containing secreted SEMA6C.1-mh or SEMA6D.1-mh were concentrated >50-fold using Centricon Plus-20 filters (Millipore; molecular mass cutoff, 10 kDa) before being used in a growth cone collapse assay. The production of SEMA6C.1-mh, SEMA6C.3-mh, SEMA6C.1-ΔPSI-mh, SEMA6D.1-mh, or SEMA6Ds-mh was monitored by Western blot with antibodies against the c-*myc* or His_6 epitope (Invitrogen). For the culture of embryonic chick E10 DRG explants, we used the method described by Goshima *et al.* (16). Primary cultures of dissociated hippocampal and cerebral cortical neurons were prepared from the brains of neonatal rats (0–1 day, Wistar) as described by Enokido *et al.* (17). Nerve growth factor-differentiated PC12 cells were cultured as described (18). The procedure for growth cone collapse assays and the method for analysis of total neurite outgrowth were used according to GrandPre *et al.* (8). For neurite outgrowth assay, PC12 cells were differentiated in the presence of NGF (100 ng/ml) for 4–7 days and then trypsinized and re-plated onto 3.5-cm wells precoated with poly-L-lysine. Simultaneously, the recombinant proteins were added to the culture. After NGF-differentiated PC12 cells were cultured for 10 to 24 h in the presence of the indicated proteins, neurite outgrowth was visualized directly or by staining with rhodamine-phalloidin.

RESULTS

Identification and Cloning of Human Semaphorin SEMA6C and SEMA6D—After sequencing the cDNA libraries of human fetal liver, a cDNA clone (No. FLD6219, 590 bp) homologous to rat *Sema6C* cDNA (13) was picked up to search the dbEST and the non-redundant GenBank™ data base. Based on a contig sequence assembly and the experimental confirmation of 5′-RACE and RT-PCR, we identified *(HSA)SEMA6C* (Fig. 1), human ortholog of rat *Sema6C* and obtained three splicing variants of SEMA6C, here named SEMA6C.1, SEMA6C.2, and SEMA6C.3, respectively. The longest isoform (SEMA6C.1, 3845 bp) contained an open reading frame capable of encoding a 962-amino acid polypeptide with a predicted molecular mass of 104.3 kDa. The translated sequence, which shows high homology (88% identity) to rat Sema6C, is composed of a sema domain (aa 64–491) followed by a PSI domain (aa 518–571) and a transmembrane segments (aa 633–653). It also contains a signal sequence (aa 1–25) at the N terminus and a proline-rich region (aa 694–874) at the C-terminal portion. The SEMA6C.2 sequence (3749 bp) encodes a 930-aa polypeptide and shares the same sequence to the SEMA6C.1 except for a 96-bp deletion between the extracellular PSI domain and the transmembrane domain. To our surprise, the SEMA6C.3 sequence (3725 bp) also shares the same sequence to the SEMA6C.1 except for a 120-bp deletion within the region coding for the sema domain; and so it encodes only a 922-aa polypeptide with an incomplete sema domain (Fig. 2*A*).

To look into whether any SEMA6C-related semaphorins re-

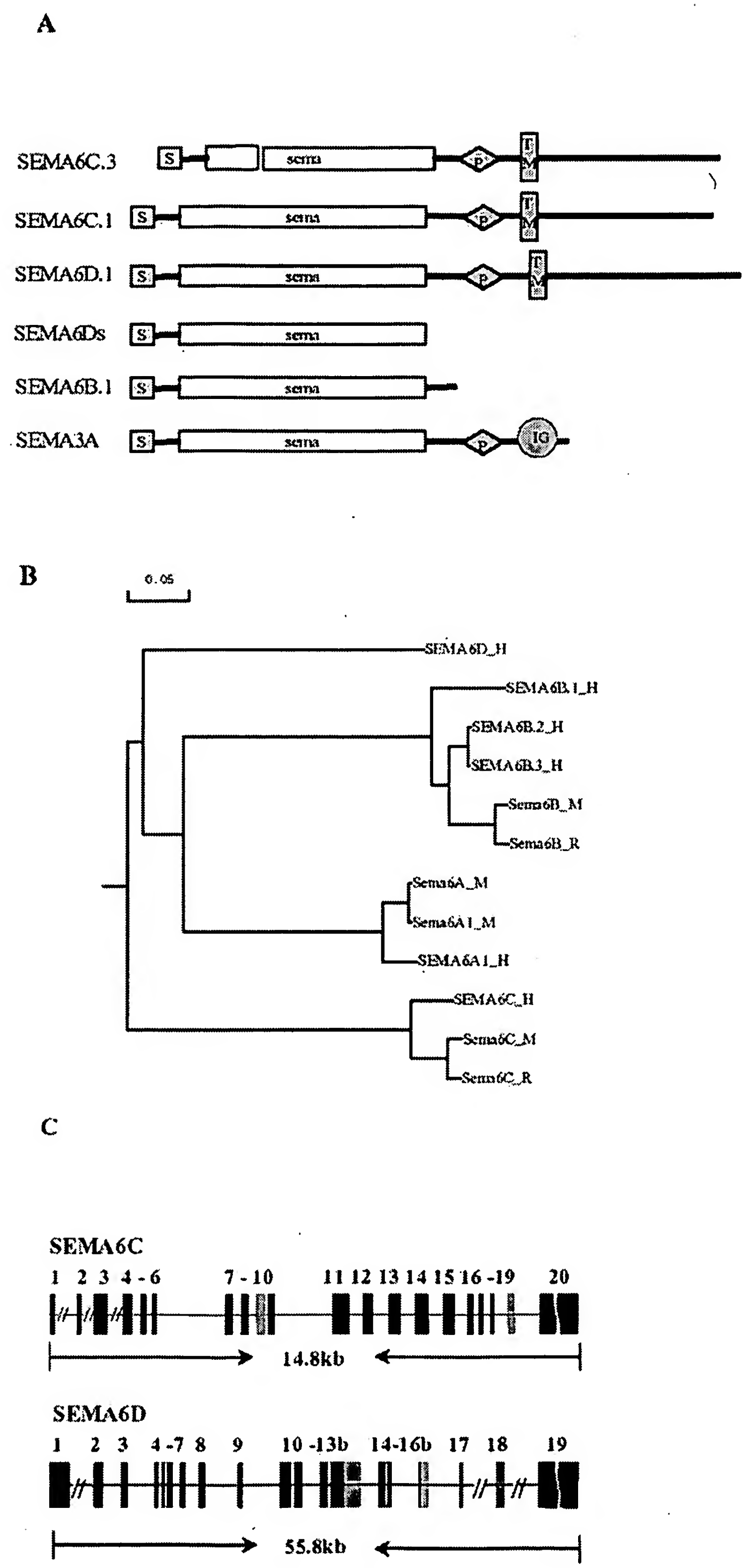

FIG. 2. **Predicted structure of (HSA)SEMA6C and (HSA)SEMA6D isoforms and genomic organization of their genes.** *A*, predicted structure of (HSA)SEMA6C and (HSA)SEMA6D isoforms are compared with related semaphorin proteins. *S*, signal sequence; *P*, PSI domain; *TM*, transmembrane segments; *IG*, Ig domain; *sema*, semaphorin domain. *B*, phylogenetic tree of mammalian class VI semaphorins is analyzed only for amino acid sequences of the sema domains and PSI domains using the CLUSTAL W program. The results were visualized by TreeView. *H*, human; *M*, mouse; *R*, rat. *C*, genomic organization of the *(HSA)SEMA6C* and *(HSA)SEMA6D* genes. *Vertical rectangles*, exons; *horizontal lines*, introns. The *shaded blocks* indicate the alternative splicing regions. For the *(HSA)SEMA6C* gene, the SEMA6C.1 isoform contains all 20 of the exons, whereas exons 9 and 19 are deleted in SEMA6C.3 and SEMA6C.2 cDNAs, respectively. For the *(HSA)SEMA6D* gene, the SEMA6Ds isoform is composed of only the first 13 exons. Each of the four long isoforms contains four to six additional exons at the 3′ region and, interestingly, uses a cryptic acceptor site in exon 13 (described as exon 13a). In detail, exons 17 and 18 are deleted in SEMA6D.1; exons 16a, 17, and 18 are deleted in SEMA6D.2; exons 16a and 18 are deleted in SEMA6D.3; and exons 16a is deleted in SEMA6D.4.

mained to be identified, we used the human *SEMA6C* sequence shown above to search against the non-redundant GenBank™ data base (BLASTn and BLASTp). Using a similar cloning strategy, we isolated five full-length cDNA sequences, which represent a single novel human gene. Amino acid sequence alignment analysis of its cDNA-coded sequence with those of (HSA)SEMA6C, rat Sema6C, and mouse Sema6C showed that it contains the typical extracellular domains of the class VI semaphorin subfamily, such as class VI sema domain and PSI domain, but differs from all of the known members of this

subfamily (Fig. 1); thus it was called (HSA)SEMA6D. Its four distinct long isoforms (here named SEMA6D.1, SEMA6D.2, SEMA6D.3, and SEMA6D.4) were 5914, 5875, 5932, and 6100 bp (not including multiple A nucleotides) in size, encoding 1011, 998, 1017, and 1073 aa, respectively. Sequence analysis has shown that all of the translated polypeptides are composed of a signal sequence (aa 1–21) followed by a class VI sema domain (aa 59–477), a PSI domain (aa 508–563), a transmembrane segment, and a long cytoplasmic region (Fig. 2*A*).

Phylogenetic analysis shows that the novel transmembrane semaphorin is closely related to class VI semaphorins, and in mammalian class VI semaphorins, semaphorin6A and -6B are closer to SEMA6D than semaphorin6C (Fig. 2*B*). In the most conserved region (the extracellular domains, *i.e.* the sema domain and PSI domain), SEMA6D not only shows considerable similarity to human SEMA6A1 (identities, 60%; positives, 77%), SEMA6C (identities, 52%; positives, 70%), and SEMA6B (identities, 52%; positives, 68%) but also moderate similarity (up to 40% amino acid identity) to many other proteins containing this domain. In addition, the PSI domain of SEMA6D contains the short cysteine-rich motif, $CX_{(8)}CX_{(5)}CX_{(3)}CX_{(7)}CX_{(7/8)}$, which is a highly conserved consensus in class VI semaphorins. Taking our findings together, we conclude that this gene is a novel member of class VI semaphorin. It was named (HSA)SEMA6D according to the views of the Semaphorin Nomenclature Committee (2).

For the four long isoforms of SEMA6D, an alternative splicing region was located between the extracellular PSI domain and the transmembrane domain(Fig. 2*C*). The shortest mRNA variant (2290 bp, here designated SEMA6Ds), however, used an early stop codon (because of a shift in the reading frame) and encoded a truncated polypeptide of 476 aa. The predicted protein is identical to the N-terminal 476-aa sequence of SEMA6D.1 and contains only a signal sequence followed by a sema domain but no PSI domain (Fig. 2*A*).

Genomic Structures and Chromosomal Localization of the Genes for Human SEMA6C and SEMA6D—When human SEMA6C and SEMA6D cDNA sequences were queried against the human genomic data base using a BLAST search, their corresponding genomic sequences were fortunately retrieved. The cDNA sequences of human SEMA6C exactly matched to chromosome 1 clone RP11-68I18 (GenBank™ accession No. AL592424.1), whereas the cDNA sequences of human SEMA6D exactly matched to chromosome 15 clone RP11-198M11 (GenBank™ accession No. AC018900.8, map 15q21.1) and chromosome 15 clone CTD-2270N23 (AC044787.6, map 15q21). Then, alignment between the human *SEMA6C* or *SEMA6D* cDNAs and their genomic sequences revealed that human SEMA6C and SEMA6D are composed of at least 20 and 19 exons, respectively (Fig. 2*C*). The sequences of the intron/exon junctions were all exactly consistent with the typical GT-AG consensus motif of the splice donor and acceptor sites except for two GC-AG sites (in introns 1 and 14) in *SEMA6C* and a GC-AG site (in intron 9) in *SEMA6D*. Alignment between the cDNAs and their genomic sequences also revealed that the human *SEMA6C* and *SEMA6D* genes span about 15 and 58 kb of genomic DNA, respectively. For human *SEMA6C*, only SEMA6C.1 contains all 20 of the exons, whereas exon 19 is deleted in SEMA6C.2 and exon 9 in SEMA6C.3. For human *SEMA6D*, no cDNA contains all 19 of the exons. Exons 13b, 17, and 18 are deleted in SEMA6D.1, whereas exons 13b, 16a, 17, and 18 are deleted in SEMA6D.2; exons 13b, 16a, and 18 are deleted in SEMA6D.3; and exons 13b and 16a are deleted in SEMA6D.4.

Following the principle of computer-based chromosomal mapping described under "Experimental Procedures," 64 ESTs highly matched to human SEMA6D cDNA sequence were collected. All of them have been clustered into a *Homo sapiens* UniGene cluster, Hs.191098, which has been mapped on chromosome 15q21.1 (www.gdb.org/gdb/) based on the sequence-tagged sites WI-6361, WI-8879, D15S1223, stSG49296, and D15S1188. Hence, the human *SEMA6D* gene should also be mapped on 15q21.1. Similarly, the human *SEMA6C* gene is assigned to chromosome 1q12–21.1.1 based on a UniGene cluster, Hs.54937 (clustered from 37 ESTs), and the marker stSG28736. The results were confirmed by the genomic sequence search described above. Recently, a search of the GenBank™ for sequence similarity with the mouse sequences turned up a high-throughput genomic sequence of a BAC (chromosome 16, clone rp23-11g21, accession No. AC084272.11) containing part of the mouse *Sema6C* gene. The sequencing of the BAC was not completed at the time of this writing.

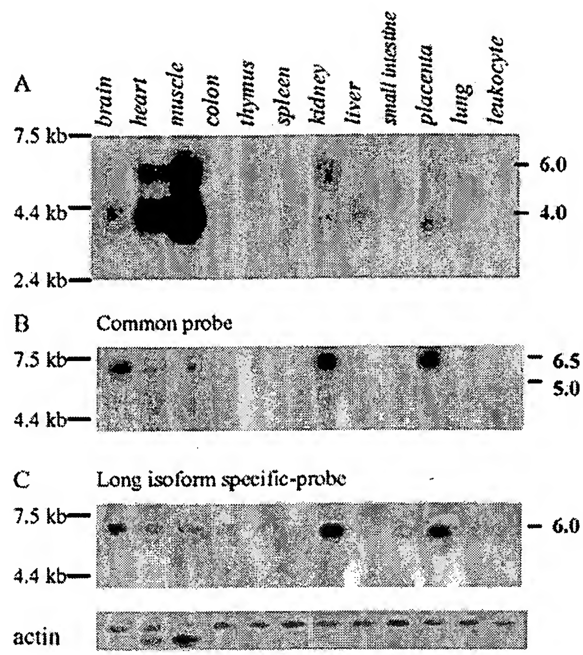

FIG. 3. **Northern blot analysis of (*HSA*)*SEMA6C* and (*HSA*)*SEMA6D* transcription in 12 human tissues.** A human multiple tissue Northern blot (1 μg of poly(A)$^+$ RNA/lane, CLONTECH) was hybridized with α-^{32}P-labeled SEMA6C-specific (*A*), SEMA6D common (*B*), or SEMA6D long isoform-specific (*C*, *upper panel*) probe or with a β-actin cDNA probe (C, *lower panel*), respectively. Hybridization with β-actin served as a loading control. Size markers are indicated on the *left*.

Expression Distribution of SEMA6C and SEMA6D—In an attempt to evaluate the transcript size and transcription profile of the human SEMA6C, a human multiple tissue Northern blot was hybridized to the SEMA6C-specific probe. Northern analysis of *SEMA6C* expression in 12 human tissues revealed two transcripts of about 4.0 and 6.0 kb (Fig. 3*A*). The smaller one (~4.0 kb) is the major transcript, the size of which is consistent with the cDNAs obtained. The larger transcript (~6.0 kb) is also long enough to cover the three isoforms of *SEMA6C* listed above. Among the 12 adult tissues, both of the transcripts are expressed predominantly in skeletal muscle, moderately in heart, brain, and kidney, and sparingly in liver and placenta, but they are hardly detectable in colon, thymus, spleen, small intestine, lung, and peripheral blood leukocytes. To evaluate the transcription profile of the short and long isoforms of human SEMA6D, Northern blot analysis of *SEMA6D* expression was carried out using probes derived from the 3′ long isoform-specific region and the 5′ common region of the SEMA6D cDNAs, respectively. A single band of ~6.5 kb

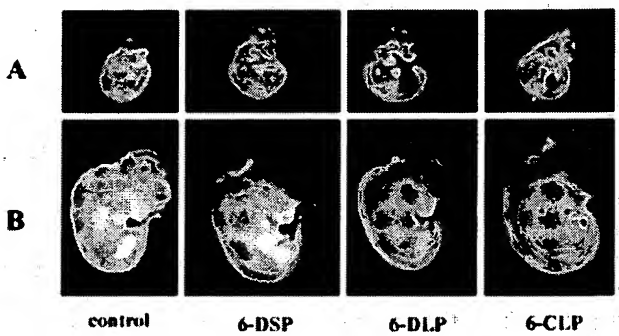

FIG. 4. ***In situ* hybridization of *(HSA)SEMA6C* and *(HSA)SEMA6D* in developmental embryos.** Murine embryos (10.5 and 13 days) were dissected free of any extra-embryonic membranes. Embryos were treated with 10 μg/ml proteinase K in PTw for 30 min and re-fixed in 4% paraformaldehyde + 0.1% glutaraldehyde for 20 min at room temperature. As much liquid as possible was removed, and the embryos were prehybridized at 63 °C for 2 h; then probe was added to a final concentration of about 0.5 μg/ml, and the embryos were hybridized overnight at 63 °C. *A*, E10.5 mouse; *B*, E13 mouse. *6-DSP*, SEMA6D short isoform probe; *6-CLP*, SEMA6C long isoform probe; and *6-DLP*, SEMA6D long isoform probe.

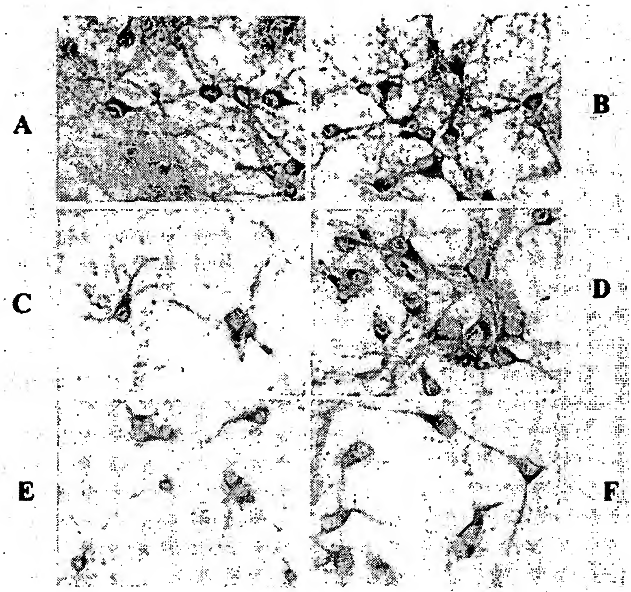

FIG. 5. ***In situ* hybridization of *(HSA)SEMA6C* and *(HSA)SEMA6D* in neurons.** *A* and *B*, photomicrographs show the neurons, after being cultured for 7 days, forming a neuronet. Dense processes were also seen between the perikaryons. All of the neurons in the figure were labeled deeply with 6-DSP (SEMA6Ds probe). The palest nuclei were noted in the center of the perikaryon; thick dendrites and deeply stained bead-like processes were labeled. No axons were labeled. *C* and *D*, the dispersion neurons and neuron cluster, after 4 days of culture, could be labeled with 6-CLP (SEMA6C long isoform probe). The soma gives rise to different numbers of deeply stained dendrites. We noted that an axon-like process seemed to be lightly stained. However, the nuclei were not stained. *E* and *F*, scattered neurons were labeled by 6-DLP (SEMA6D long isoform probe) after culture for 4 days. The perikaryons were labeled lightly or deeply, but there was no labeling in the nuclei.

was detected with the 3′-specific probe in most of the 12 tissues (Fig. 3*C*), whereas two distinct transcripts (6.5 and 5.0 kb) were identified with the 5′ common probe (Fig. 3*B*). The larger one, the size of which is consistent with the long isoforms of *SEMA6D*, is the major transcript. It is expressed abundantly in kidney, brain, and placenta, moderately in heart and skeletal muscle, and sparingly in the lung, colon, and small intestine but is hardly detectable in liver, spleen, thymus, and peripheral blood leukocytes. The smaller transcript was expressed faintly in kidney, skeletal muscle, heart, and placenta and was hardly detectable in the other eight tissues. The results presented above were confirmed by an additional dot blot analysis, in which a dot blot containing a total of 68 normal tissues and 8 human cancer lines was hybridized to the same probes for *SEMA6C* or *SEMA6D*, respectively. Intriguingly, *SEMA6C* or *SEMA6Ds* was hardly detectable in the eight human cancer cell lines examined (detailed data are shown in the Supplemental Material).

In order to get further insight into the role(s) played by these semaphorins during neural development, we carried out *in situ* hybridization of murine whole mount embryos. It was demonstrated that there are no hybridization signals in E10.5 mouse. However, for E13 mouse embryos, hybridization staining was profound in the brains of all experimental groups. The distribution of positive signals (Fig. 4) was located mainly around the areas of the lateral ventricle, striatum, wall of midbrain, pons/midbrain junction, and choroid plexus for 6-DSP (SEMA6Ds probe), 6-CLP (SEMA6C long isoform probe), and 6-DLP (SEMA6D long isoform probe). Meanwhile, staining on the roof of the neopallial cortex was also found in 6-DLP and on the spinal cord in 6-CLP. The results presented above imply important roles for the three genes during the development of the central nervous system, but there is also little difference among them.

Moreover, the expression of these semaphorins in neurons was identified. The primary neurons, isolated from the spinal cords (cortex) of newborn rats, showed a typical appearance of different kinds of neurons, with different sizes and shapes. Under a phase contrast microscope, they became hypertrophic and enlarged beginning at 6 h, and then different numbers of thin processes were seen to set at each of its termini or around the soma (see Supplemental Material). A few glias were also noted in the dishes. Two days later, the neuron in the dishes formed a net, and more dense processes also formed. Fig. 5 shows that almost all of the neurons in *in situ* hybridization dishes were stained by three kinds of special probes such as 6-DSP (SEMA6Ds probe), 6-CLP (SEMA6C long isoform probe), and 6-DLP (SEMA6D long isoform probe). However, no astrocytes in the dishes were stained. The labeled neurons, including their perikaryons and processes and even their termini, were stained with a deep blue color; however, their nuclei were not stained. We carefully observed the axons but did not find any labeling. There was no labeling in the control group or in the perikaryons and processes.

Alternative Splicing of SEMA6C and SEMA6D—As three SEMA6C alternatively spliced variants and four SEMA6D long isoforms were isolated, it was implied that they might have different expression patterns. To evaluate this possibility for *SEMA6C*, we first performed RT-PCR assays using primers designed to give 453- or 357-bp bands depending on the presence or absence of the 96-bp insertion. As shown in Fig. 6, mRNA from most tissues gave both the 453-bp band and the 357-band; however, in muscle the 453-bp band (corresponding to SEMA6C.1 and SEMA6C.3) was far stronger than the 357-band (corresponding to SEMA6C.2), which is consistent with the result from a previous report on the rat ortholog (13). When human cerebral cortex and cerebellum were examined by the same analysis, however, only the 357-bp band was detected, indicating that SEMA6C.2 is the major transcript in these tissues. To verify the existence of the human SEMA6C.3 sequence above and assess the expression profile of the distinct isoforms, we performed RT-PCR using primers designed to give 495- or 375-bp bands depending on the presence or absence of the 120-bp insertion. In most of the tissues tested (*i.e.* human

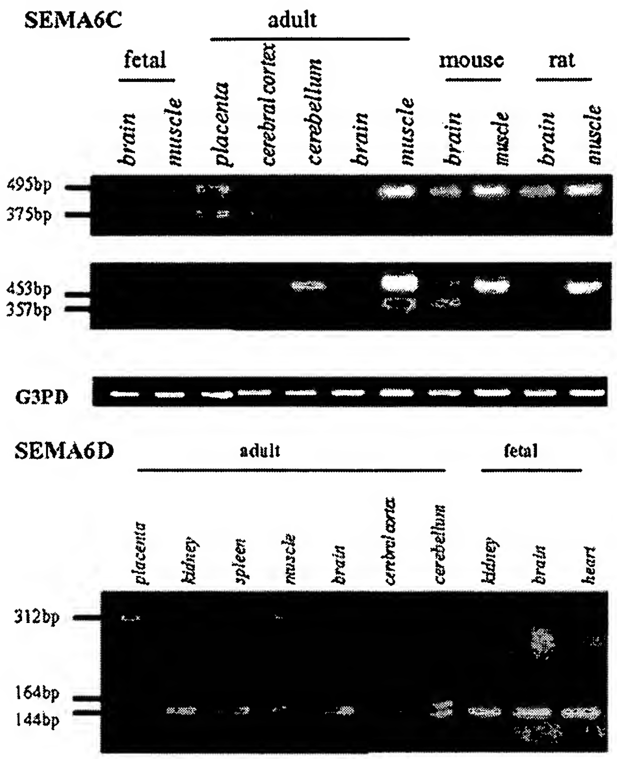

FIG. 6. **Distribution of (HSA)SEMA6C and (HSA)SEMA6D isoforms examined by RT-PCR.** Human mRNA (CLONTECH) or rat and mouse total RNA from various tissues was used as template for RT-PCR. For (HSA)SEMA6C (*top*), a pair of primers gave 495- and 375-bp fragments, indicating the presence or deletion of a 120-bp insertion; another pair of primers gave 453- and 357-bp fragments indicating the presence or deletion of a 96-bp insertion. Panels were normalized against glyceraldehyde-3-phosphate dehydrogenase. For (HSA)SEMA6D (*bottom*), a pair of primers gave 144-, 164-, and 312-bp fragments representing SEMA6D.1, SEMA6D.3, and SEMA6D.4, respectively. The predicted 105-bp band (representing SEMA6D.2) was not detected in any tissue examined so far. Amplified DNA fragments were fractionated on a 1% (SEMA6C) or 2% (SEMA6D) agarose gel and stained by ethidium bromide. The fragment sizes are shown on the *left*.

placenta and cerebellum and human, mouse, and rat adult skeletal muscle), both the 495-bp band (corresponding to SEMA6C.1 and SEMA6C.2) and the 375-bp band (corresponding to SEMA6C.3) were detected, indicating ubiquitous expression of the isoforms. In muscle, however, the 495-bp band was much stronger than the 375-band. Thus, the results presented above suggest that SEMA6C.1 is the major isoform of SEMA6C. In addition, almost none of the 375-bp band was detected in the adult human cerebral cortex, suggesting that SEMA6C.3 is expressed exclusively in certain tissues.

As described above (Fig. 2*C*), human SEMA6D mRNAs include one short isoform and four long isoforms. The SEMA6D.2 isoform (5875 bp) shares the same sequence to SEMA6D.1 except for a 39-bp deletion between the extracellular PSI domain and the transmembrane domain. The SEMA6D.3 isoform (5932 bp) shares the same sequence to the SEMA6D.2 sequence except for a 57-bp insertion, whereas the 6100 bp isoform (SEMA6D.4) shares the same sequence to the SEMA6D.3 sequence except for a 168-bp insertion nearby the above region. To further verify the existence of the four long isoforms for human SEMA6D and assess their relative expression profile, we performed RT-PCR using primers designed to give 105-, 144-, 162-, or 312-bp bands corresponding to SEMA6D.2, SEMA6D.1, SEMA6D.3, and SEMA6D.4, respectively. As shown in Fig. 6, the 144 bp band was ubiquitously detected in the adult and fetal human tissues examined except for the placenta, only the 162-band in almost all the adult tissues, and a significant 312-bp band only in adult muscle and placenta, indicating that SEMA6D.1 is the major transcript. However, the predicted 105-bp band was not detected in any examined tissue so far.

Secretion and Cellular Localization of SEMA6C and SEMA6D—Hydrophilic analysis shows that each of the three isoforms of SEMA6C or the four long isoforms of SEMA6D contains two hydrophobic regions. One is assumed to be a signal sequence and the other a transmembrane sequence. Thus, all of the proteins are assumed to be cell surface proteins. On the contrary, since SEMA6Ds contains only a signal sequence followed by a sema domain (Fig. 2*A*), it is assumed to be a secreted protein. Therefore, when their partial cDNAs encoding the extracellular regions (including a signal sequence, a sema domain, and a PSI domain) or the entire CDS of SEMA6Ds were subcloned into pcDNA3.1-MycHis and expressed transiently in COS7 cells, it was predicted that the recombinant proteins would be expressed in soluble form. To confirm this prediction and to facilitate functional studies of the SEMA6C and SEMA6D genes, the four plasmids pcDNA3.1-SEMA6C.1, pcDNA3.1-SEMA6C.3, pcDNA3.1-SEMA6Ds, and pcDNA3.1-SEMA6D.1, described under "Experimental Procedures," were expressed transiently in COS7 cells. Expression of the proteins was confirmed by Western analysis. A unique band is specifically recognized in each of the cell extracts or the supernatants collected from pcDNA3.1-SEMA6C.1- or pcDNA3.1-SEMA6D.1-transfected cells (Fig. 7*A*). In the cell lysates, the identified bands are consistent with the predicted sizes of 62.5 and 67.1 kDa for the unprocessed and the tagged fusion proteins (including the signal peptide), respectively. In the supernatants, however, the apparent molecular masses (~75 kDa and 80 kDa, respectively) were larger than the predicted sizes of 59.7 and 64.7 kDa (the processed fusion proteins without the signal peptide). This suggests that the secreted proteins were modified during or after secretion. Rat Sema6C and other semaphorins have been demonstrated as glycoproteins (3, 13), and so it is very likely that SEMA6C and SEMA6D are also glycoproteins. This probability was supported by the motif searching through Prosite, which indicated that there are three potential *N*-glycosylation sites in SEMA6C.1 and nine in SEMA6D.1.

In the cell extracts collected from pcDNA3.1, SEMA6C.3-, or pcDNA3.1-SEMA6Ds-transfected cells, the unique bands are also specifically recognized with the predicted sizes of 58.1 and 54.2 kDa for the unprocessed and the tagged fusion proteins, respectively (Fig. 7*B*). In the conditioned medium, however, no band was specifically recognized, indicating that the fusion proteins were not secreted. Based on the molecular masses of the fusion proteins in the cell lysates, the fusion proteins within the host cells might not be modified with *N*-linked glycosylation. To evaluate further the possible roles of different extracellular domains on subcellular localization of the semaphorin proteins, we constructed a vector (pcDNA3.1-SEMA6C.1-ΔPSI) by deleting the PSI domain of pcDNA3.1-SEMA6C.1. COS7 cells transfected with this vector did not express the fusion protein in secreted form either, which is consistent to the results above. To examine further whether the incomplete extracellular domains have any effect on the intracellular localization of the proteins, we performed immunostaining of COS7 cells expressing the five fusion proteins described above using anti-Myc antibody. As predicted, COS7 cells transfected with pcDNA3.1-SEMA6C.1 or pcDNA3.1-SEMA6D.1 showed Myc immunoreactivity on the cell surface and in the cytoplasm (Fig. 7*C*). In the cells transfected

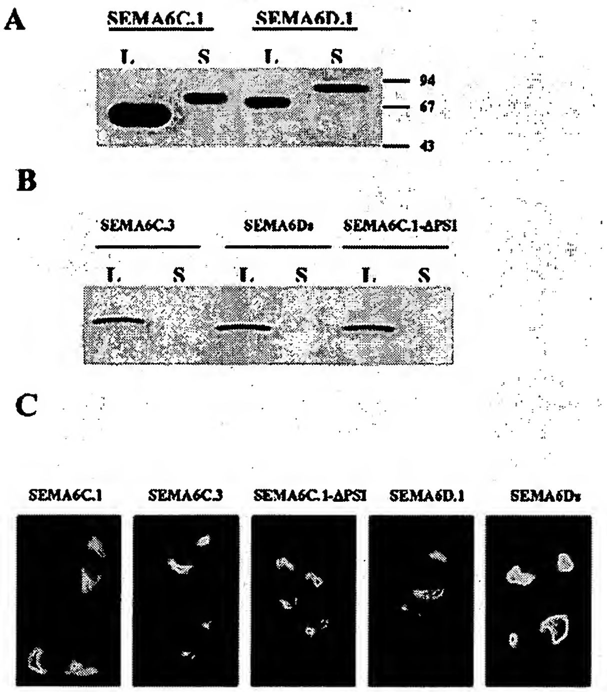

FIG. 7. **Characterization of SEMA6C and SEMA6D transiently expressed in COS7 cells.** The recombinant proteins (SEMA6C.1-mh, SEMA6C.3-mh, SEMA6C.1-ΔPSI-mh, SEMA6D.1-mh, and SEMA6Ds-mh) from both the cell lysates and the conditioned media (supernatants) were run on a 10% SDS-polyacrylamide gel under reducing condition and visualized on a Western blot with a monoclonal antibody against the Myc tag. The predicted bands are recognized specifically in both the cell lysates (*L*) and the supernatants (*S*) collected from pcDNA3.1-SEMA6C.1- or pcDNA3.1-SEMA6D.1-transfected cells (*A*). The 62- and 67-kDa protein bands from the cell lysates containing SEMA6C.1-mh and SEMA6D.1-mh are consistent with the predicted sizes of the unprocessed and the tagged proteins, respectively. The predicted fusion proteins were detected also in the cell lysates from transfected cells with pcDNA3.1-SEMA6C.3, pcDNA3.1-SEMA6C.1-ΔPSI, or pcDNA3.1-SEMA6Ds but not in the corresponding conditioned media (*B*). *C*, cellular localization of the five fusion proteins displayed by immunostaining of transfected COS7 cells. COS7 cells were transfected with the five plasmids listed above. Two days after the transfection, the cells were fixed with paraformaldehyde and stained with alkaline phosphatase-labeled anti-Myc antibody. The nuclei were stained by Hoechst 33342 (data not shown).

with pcDNA3.1-SEMA6C.3, pcDNA3.1-SEMA6C.1-ΔPSI, or pcDNA3.1-SEMA6Ds, the fusion protein was detected as punctate green staining distributed throughout the cytoplasm and even accumulated at the two ends of many cells, which counted 56.8, 45.2, and 62.5% of the total transfected cells, respectively. In the other cells, however, the distribution of the immunoreactive material was similar to that in the cells transfected with pcDNA3.1-SEMA6C.1 or pcDNA3.1-SEMA6D.1. As seen in Fig. 2*A*, the expected peptide encoded by pcDNA3.1-SEMA6C.3 contained a truncated sema domain, and the peptide encoded by pcDNA3.1-SEMA6C.1-ΔPSI or pcDNA3.1-SEMA6Ds contained no PSI domain at all. Thus, the results support the possibility that a sema domain and a PSI domain in their integrity are necessary for the appropriate post-translational modification and subcellular localization of the semaphorin proteins. Those results were further supported by analysis with the GFP fusion protein localization system (see Supplemental Materials).

Growth Cone Collapse Activity of SEMA6C and SEMA6D—Because SEMA6C.1 and SEMA6D.1 are demonstrated to be the major isoforms of SEMA6C and SEMA6D, respectively, soluble versions of SEMA6C.1 and SEMA6D.1 were engineered for the functional experiments. Medium conditioned by the cells transfected with SEMA6C.1 or SEMA6D.1 was concentrated and added to chick E10 DRG explant and cultures, respectively. Growth cone morphology was assessed after a 60-min incubation at 37 °C by fixation and rhodamine-phalloidin staining. The extracellular domain of SEMA6C.1 possesses growth cone-collapsing activity for chick E10 DRG neurons, acutely altering growth cone morphology at concentrations as low as 1.0 mg/ml. It is consistent with the results of rat Sema6C (13). In addition, the supernatant containing SEMA6C.1-mh collapsed growth cones of cultured rat hippocampal neurons and rat cortical neurons in a dose-responsive manner (Fig. 8, *A* and *B*). In comparison, SEMA6D.1-mh was also observed to inhibit axonal extension of NGF-differentiated PC12 cells and to collapse growth cones of chick DRG and rat hippocampal neurons but not to collapse growth cones of rat cortical neurons, even at concentrations as high as 4 mg/ml (Fig. 8*C*). As a control, medium conditioned by mock-transfected cells had no apparent collapsing activity. The growth cone collapse and outgrowth assays suggest that SEMA6C.1 and SEMA6D.1 inhibit axon outgrowth activity with distinct tissue specificity.

DISCUSSION

So far more than 25 semaphorin genes have been cloned, and some of them have been shown to be repulsive signals for growing axons. In the class VI semaphorins, only mouse Sema6A were reported to have growth cone collapse activities on chick E8 sympathetic chain ganglia and E7 DRG (19); rat Sema6C has similar activity on chick E8 DRG (13). In this study, to determine whether SEMA6C and SEMA6D have repulsive activity on neurons, we performed growth cone collapse assays and a neurite outgrowth assay using their recombinant secreted proteins instead of the wild-type transmembrane proteins. Functionally, the extracellular domain of SEMA6C inhibited axonal extension of NGF-differentiated PC12 cells and induced growth cone collapse of chick DRG, rat hippocampal neurons, and rat cortical neurons in a dose-dependent manner. SEMA6D was also observed to inhibit axonal extension of NGF-differentiated PC12 cells and collapse growth cones of chick DRG and rat hippocampal neurons but not to collapse growth cones of rat cortical neurons at the tested concentrations. Under our experimental conditions, it seems that SEMA6C and SEMA6D have similar potent repulsive activity on the neurons examined, except on rat cortical neurons. These data suggest that membrane-bound semaphorins act as chemo-inhibitory molecules in neuronal development, a role similar to and likely overlapping with that of the class III semaphorins. However, the expression of SEMA6C and SEMA6D predominantly in the adult tissues makes them potentially important molecules in nervous system maintenance and repair. Furthermore, the sensitivities of different growth cones to distinct semaphorins may be different.

It is worth noting that SEMA6D has an expression profile that overlaps that of SEMA6A-1, a very close relative of Sema6A. SEMA6A-1 is expressed highly in the placenta and fetal brain and kidney (20), whereas SEMA6D is expressed abundantly not only in the placenta and fetal brain and kidney but also in the adult brain and kidney. This profile is consistent with a more general role of the proteins in neurogenesis and organogenesis as well as in regenerative and degenerative processes; all of these expression areas are characterized by a highly dynamic rearrangement of cytoskeletal elements (20). It has been reported that Sema6A has an expression pattern consistent with a role as a local inhibitor of developing sympathetic axons *in vivo*. In addition, another "function unknown" semaphorin, SEMA6B, was expressed strongly in brain and moderately in heart (21). Therefore, it is likely that many of

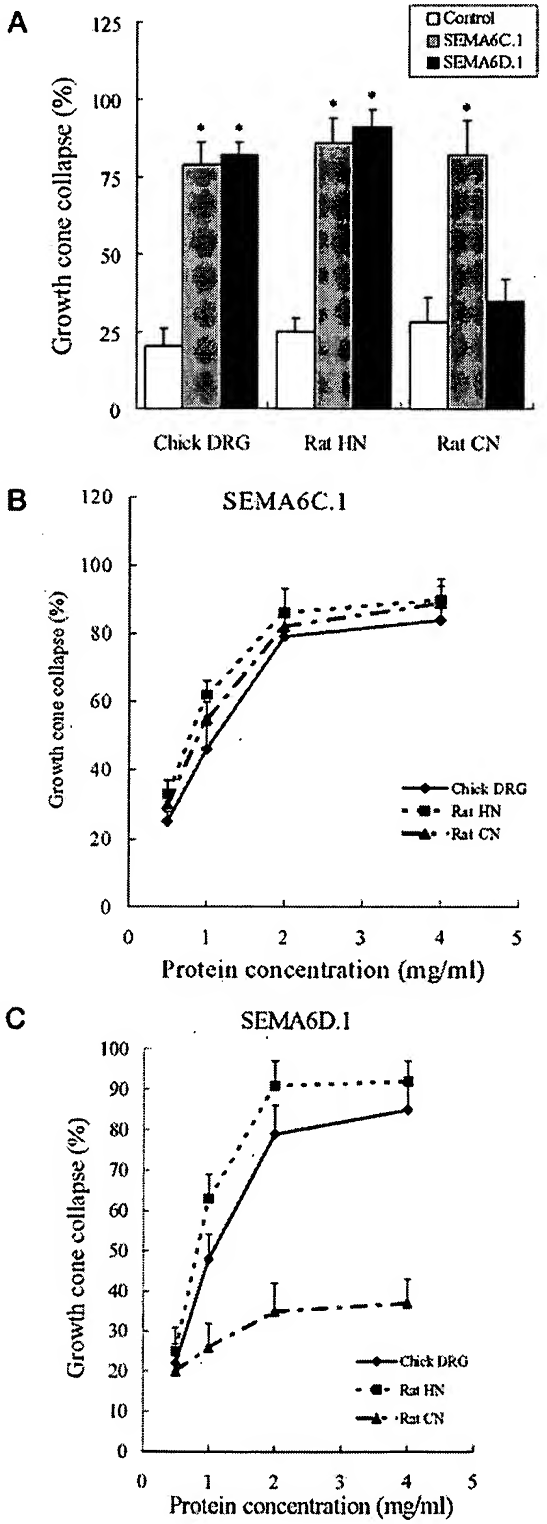

FIG. 8. **The inhibitory activities of SEMA6C-mh and SEMA6D-mh for axonal extension.** Chick E10 explants and dissociated rat hippocampal (*HN*) and cortical neurons (*CN*) were cultured and exposed to SEMA6C-mh and SEMA6D-mh collected from the conditioned media for 60 min before fixation and staining with rhodamine-phalloidin. The percentage of collapsed growth cones at 2 mg of total protein/ml (*A*) and the dose-response curves comparing the growth cone collapsing activities of SEMA6C-mh (*B*) and SEMA6D-mh (*C*) on chick E10 explants and dissociated rat HN and CN are shown. Conditioned medium containing mock MycHis was used as a negative control. All results are the means ± S.E. from four to six determinations. In *A* and *D*, those values that are significantly different from control are indicated (*, $p < 0.05$).

these overlapping molecules have redundant functions acting on any class of axons, but their probable precise combinations guiding many classes of axons, by process of elimination, to their target tissues need to be revealed.

Because three isoforms of SEMA6C and five isoforms of SEMA6D have been isolated, the significance of the alternatively spliced variants attracted our interest. With Northern blot analysis and RT-PCR, we have demonstrated their expression to be regulated in a tissue- and development-dependent manner; SEMA6C.1 and SEMA6D.1 are the major isoforms of SEMA6C and SEMA6D, respectively. Similarly to rat Sema6C (13), the alternatively splicing region is located mainly between the extracellular PSI domain and transmembrane segments for human SEMA6C and SEMA6D (*i.e.* SEMA6C.2 and the four long isoforms of SEMA6D). What is of interest is that SEMA6C.3 contains only an incomplete sema domain, and SEMA6Ds is a "SEMA6D truncate" lacking the PSI domain. The alternative splicing process related to deletion in the sema domain was described previously in human SEMA3F (22, 23) and SEMA4F (24). The biological significance of these deletions in the sema domain is inexplicable, because the locations of those deletions correspond to the position of an important 70 amino acid region within the sema domain, which was shown to specify the biological activity of the three class III semaphorins by deletion analysis (25). In addition, similar to SEMA6Ds, a human "semaphorin 6B truncate" (SEMA6B.1) containing a signal sequence followed only by a sema domain but lacking the PSI domain has also been assumed to be a secreted protein by others (21). When the entire CDS of SEMA6Ds was subcloned into pcDNA3.1-MycHis and expressed transiently in COS7 cells, however, the recombinant protein was present only in the cell lysate but was not detected in the conditioned medium. SEMA6B.1 has not been confirmed to represent a secreted protein by any experiment either. Intriguingly, all of the COS7 cells transfected with the three expression vectors (pcDNA3.1-SEMA6C.3, pcDNA3.1-SEMA6C.1-ΔPSI, or pcDNA3.1-SEMA6Ds) encoding incomplete extracellular domains could not secrete the recombinant proteins into the medium. These data plus the immunofluorescence assay suggest that the integrity of a sema domain or a PSI domain is necessary for the correct subcellular localization of the semaphorin proteins. Because the extracellular region of SEMA4D (CD100) has recently been demonstrated to be released from the surface of T lymphocytes by regulated proteolysis and thus to act as a long range guidance cue in the immune system (26), it is worthwhile to study whether this kind of post-translational modification is shared by other transmembrane semaphorins.

Although the cytoplasmic regions of class VI semaphorins are longer than those of other classes, which are sufficient to have signaling function (20, 24, 27, 28), recent work has begun in elucidating the nature of the semaphorin receptor. So far, transmembrane semaphorins other than class III have been shown to interact directly with plexins (29–31). Because the plexin family is a large one, it is attractive to suppose that particular plexins will help to determine the specificity of both semaphorin binding and the biological response. It is hoped that the characterization of the receptors for SEMA6C and SEMA6D and the further elucidation of their functions *in vivo* will facilitate the understanding of the mechanism of neural development and nerve regeneration after injury and provide us with possible treatment strategies for certain neurodegenerative diseases.

REFERENCES

1. Tamagnone, L., and Comoglio, P. M. (2000) *Trends Cell Biol.* **10,** 377–383
2. Semaphorin Nomenclature Committee (1999) *Cell* **97,** 551–552
3. Raper, J. A. (2000) *Curr. Opin. Neurobiol.* **10,** 88–94
4. Serafini, T., Kennedy, T. E., Galko, M. J., Mirzayan, C., Jessell, T. M., and Tessier-Lavigne, M. (1994) *Cell* **78,** 409–423
5. Holder, N., and Klein, R. (1999) *Development* **126,** 2033–2044
6. Kramer, S. G., Kidd, T., Simpson, J. H., and Goodman, C. S. (2001) *Science* **292,** 737–740
7. Wu, W., Wong, K., Chen, J., Jiang, Z., Dupuis, S., Wu, J. Y., and Rao, Y. (1999) *Nature* **400,** 331–336
8. GrandPre, T., Nakamura, F., Vartanian, T., and Strittmatter, S. M. (2000) *Nature* **403,** 439–444
9. Bork, P., Doerks, T., Springer, T. A., and Snel, B. (1999) *Trends Biochem. Sci.* **24,** 261–263
10. Wang, G., Yang, X. M., Zhang, Y., Wang, Q. M., Chen, H. P., Wei, H. D., Xing, G. C., Xie, L., Zhang, C. G., Fang, D. C., Wu, C. T., and He, F. C. (1999) *J. Biol. Chem.* **274,** 11469–11472
11. Qu, X., Zhang, C., Zhai, Y., Xing, G., Wei, H., Yu, Y., Wu, S., and He, F. (2001) *Gene* **264,** 37–44
12. Yu, Y., Zhang, C., Zhou, G., Wu, S., Qu, X., Wei, H., Xing, G., Zhai, Y., Wan, J., Ouyang, S., Li, L., Zhang S., Wu, C., and He, F. (2001) *Genome Res.* **11,** 1392–1403
13. Kikuchi, K., Chedotal, A., Hanafusa, H., Ujimasa, Y., de Castro, F., Goodman, C. S., and Kimura, T. (1999) *Mol Cell Neurosci.* **13,** 9–23
14. Harlow, E., and Lane, D. (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
15. Luo, Y., Raible, D., and Raper, J. A. (1993) *Cell* **75,** 217–227
16. Goshima, Y., Nakamura, F., Strittmatter, P., and Strittmatter, S. M. (1995) *Nature* **376,** 509–514
17. Enokido, Y., Akaneya, Y., Niinobe, M., Mikoshiba, K., and Hatanaka, H. (1992) *Brain Res.* **599,** 261–271
18. Strittmatter, S. M., Fishman, M. C., and Zhu, X. P. (1994) *J. Neurosci.* **14,** 2327–2338
19. Xu, X. M., Fisher, D. A., Zhou, L., White, F. A., Ng, S., Snider, W. D., and Luo, Y. (2000) *J. Neurosci.* **20,** 2638–2648
20. Klostermann, A., Lutz, B., Gertler, F., and Behl, C. (2000) *J. Biol. Chem.* **275,** 39647–39653
21. Correa, R. G., Sasahara, R. M., Bengtson, M. H., Katayama, M. L., Salim, A. C., Brentani, M. M., Sogayar, M. C., de Souza, S. J., and Simpson, A. J. (2001) *Genomics* **73,** 343–348
22. Roche, J., Boldog, F., Robinson, M., Robinson, L., Varella-Garcia, M., Swanton, M., Waggoner, B., Fishel, R., Franklin, W., Gemmill, R., and Drabkin, H. (1996) *Oncogene* **12,** 1289–1297
23. Sekido, Y., Bader, S., Latif, F., Chen, J. Y., Duh, F. M., Wei, M. H., Albanesi, J. P., Lee, C. C., Lerman, M. I., and Minna, J. D. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 4120–4125
24. Encinas, J. A., Kikuchi, K., Chedotal, A., de Castro, F., Goodman, C. S., and Kimura, T. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 2491–2496
25. Koppel, A. M., Feiner, L., Kobayashi, H., and Raper, J. A. (1997) *Neuron* **19,** 531–537
26. Delaire, S., Billard, C., Tordjman, R., Chedotal, A., Elhabazi, A., Bensussan, A., and Boumsell, L. (2001) *J. Immunol.* **166,** 4348–4354
27. Inagaki, S., Ohoka, Y., Sugimoto, H., Fujioka, S., Amazaki, M., Kurinami, H., Miyazaki, N., Tohyama, M., and Furuyama, T. (2001) *J. Biol. Chem.* **276,** 9174–9181
28. Wang, L. H., Kalb, R. G., and Strittmatter, S. M. (1999) *J. Biol. Chem.* **274,** 14137–14146
29. Takahashi, T., Fournier, A., Nakamura, F., Wang, L. H., Murakami, Y., Kalb, R. G., Fujisawa, H., and Strittmatter, S. M. (1999) *Cell* **99,** 59–69
30. Tamagnone, L., Artigiani, S., Chen, H., He, Z., Ming, GI., Song, H., Chedotal, A., Winberg, M. L., Goodman, C. S., Poo, M., Tessier-Lavigne, M., and Comoglio, P. M. (1999) *Cell* **99,** 71–80
31. Vikis, H. G., Li, W., He, Z., and Guan, K. L. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 12457–12462